

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-217961

(43)Date of publication of application : 02.08.2002

(51)Int.Cl.

H04L 12/56

H04L 12/40

(21)Application number : 2001-015615

(71)Applicant : FUJITSU LTD

(22)Date of filing : 24.01.2001

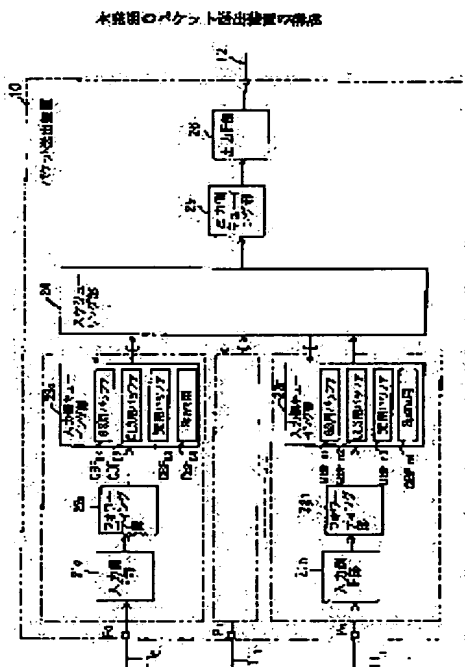
(72)Inventor : YAMAMOTO KANTA

## (54) PACKET-TRANSMITTING DEVICE

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To enable band control, without having to divide a variable-length packet.

**SOLUTION:** A packet-transmitting device is provided, which transmits packets belonging to a plurality of groups with different priorities. Here, a queue control part of a scheduling part 24 generates a queue for each group, and sequentially provides, for each queue, the component of the queue with a packet transmission right by a round robin method. A transmitted-group deciding part of the scheduling part 24 decides a group of top priority as a packet transmission group among groups, in which at least one packet corresponding to a queue component is waiting to be transmitted. A queue control part gives packet transmission right to a top-priority component of the queue in the decided transmission group, and transmits a packet corresponding to the component.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

BEST AVAILABLE COPY

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-217961

(P2002-217961A)

(43) 公開日 平成14年8月2日(2002.8.2)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	ターム(参考)
H 0 4 L 12/56	2 0 0	H 0 4 L 12/56	2 0 0 C 5 K 0 3 0
12/40		12/40	D 5 K 0 3 2

審査請求 未請求 請求項の数 5 O L (全 17 頁)

(21) 出願番号 特願2001-15615(P2001-15615)

(22) 出願日 平成13年1月24日(2001.1.24)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番  
1号

(72) 発明者 山本 幹太

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

(74) 代理人 100084711

弁理士 斉藤 千幹

Fターム(参考) 5K030 HA08 HB17 HD03 HD06 KA03

5K032 CA17 CD01 DB20 DB26

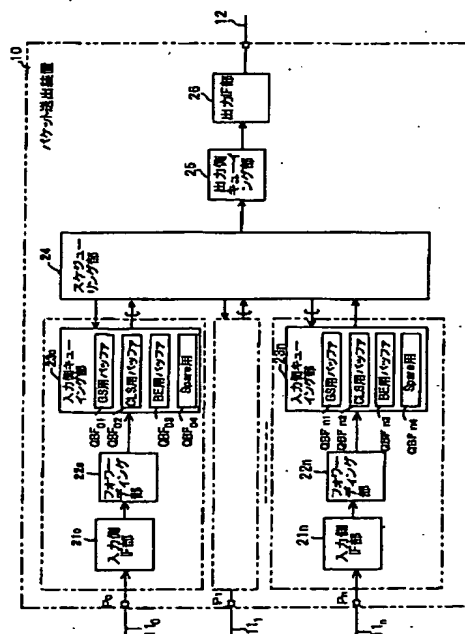
(54) 【発明の名称】 バケット送出装置

(57) 【要約】

【課題】 可変長バケットを分割することなく帯域制御を可能にする。

【解決手段】 優先度がそれぞれ異なる複数のグループに属するバケットを送出するバケット送出装置である。このバケット送出装置において、スケジューリング部24の待ち行列制御部はグループ毎に待ち行列を生成し、待ち行列毎に該待ち行列を構成する要素にラウンドロビン方式に従って順番にバケット送出権を与える。スケジューリング部24の送出グループ決定部は、少なくとも1つの待ち行列構成要素に応じたバケットが送出待ちになっているグループのうち、優先度が最大のグループをバケット送出グループとして決定し、待ち行列制御部は決定された送出グループの待ち行列の最優先行列構成要素にバケット送出権を与え、該行列構成要素に応じたバケットを送出する。

本発明のバケット送出装置の構成



## 【特許請求の範囲】

【請求項1】 優先度がそれぞれ異なる複数のグループに属するバケットを順に送出するバケット送出装置において、

グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にバケット送出権を与える待ち行列制御部、少なくとも1つの待ち行列構成要素に応じたバケットが送出待ちになっているグループのうち、優先度が最大のグループをバケット送出グループとして決定するバケット送出グループ決定部、

を備え、待ち行列制御部は前記決定されたバケット送出グループの待ち行列においてバケット送出権のある行列構成要素に応じたバケットを送出することを特徴とするバケット送出装置。

【請求項2】 入力ポートと該ポートより入力するバケットに付加する各品質クラスの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部、を備え、前記待ち行列制御部は、待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にバケット送出権を与える、ことを特徴とする請求項1記載のバケット送出装置。

【請求項3】 待ち行列構成要素毎に送出待ちのバケットを記憶するバッファ部、送出待ちバケットが記憶されているバッファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部、を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちバケットが存在するグループを識別し、これらグループのうち優先度が最大のグループをバケット送出グループとして決定する、ことを特徴とする請求項1記載のバケット送出装置。

【請求項4】 行列構成要素毎に1回のバケット送出権で保証するデータ送出量を設定する保証データ量設定部、バケット送出権を与えられた行列構成要素に応じたバケットの実際の送出量を監視する監視部、実際のデータ送出量が前記保証データ量と等しくなったとき、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部、を備え、前記待ち行列制御部は、該制御信号に基づいてバケット送出権を次の行列構成要素に付与する、ことを特徴とする請求項1または2記載のバケット送出装置。

【請求項5】 待ち行列構成要素毎に設定時間当りに送出するバケットのデータ送出流量を設定するデータ送出流量設定部、待ち行列構成要素毎に前記設定時間当りのバケットの実際のデータ送出流量を監視する手段、

待ち行列構成要素毎にバケットの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたバケットの送出を禁止する送出禁止信号を発生するバケット送出禁止手段、を備えたことを特徴とする請求項1又は2記載のバケット送出装置。

## 【発明の詳細な説明】

【0001】

10 【発明の属する技術分野】本発明は優先度がそれぞれ異なる複数のグループに属するバケットを順に送出するバケット送出装置に係わり、特に、可変長バケットのスイッチングを行うIPルータ装置、Ethernetスイッチ装置などのバケット送出装置に関する。

【0002】

【従来の技術】IPネットワーク(インターネット)が登場した頃、ネットワーク内では差別化処理ではなく、ある特定のフローに帯域を優先して割り当てるなどの処理は無かった。この初期のインターネット形態では、帯域の割り当ては「早い者勝ち」の世界であり、多くデータを送れば送るほどネットワークの帯域を占有するというもので、スケジューリング論理も単純なFiFo(First In First Out)であった。しかし、時代は次第にIPネットワークに、電子ファイル・データ、e-mail、静止画像だけではなく、動画、音声というマルチメディア情報を送れる事を期待しはじめた。動画、音声というマルチメディア情報は、いわゆる「リアルタイム・トラフィック」と呼ばれるように、送信元から送信先に対して、「ある一定帯域、一定周期で、データの損失無し」で、通信される事が期待されているトラフィックである。なお、これまでのリアルタイム・トラフィックは、送信元から送信先への固定的な帯域を確保してから通信する、いわゆる電話網にてやり取りされていたトラフィックであった。

30 【0003】また、急激に増加したインターネット・ユーザからの膨大なバケット・データは、ネットワーク内でバケットの混雑状態を招き、多くのバケットがネットワーク内で処理できず廃棄され始めた。そうすると、多くのユーザにおいて、目に見て分かるくらいのthroughputの低下が感じられるようになった。ここに、「私は人より多くお金を払うので、throughputを上げてくれ」という要望が生まれる。これはよく、車の一般道と高速道路に例えられる。固定的な帯域を保証してもらうのは高速道路、いろんなトラフィックが帯域の奪い合いをやっているのが一般道、という事だが、注意が必要なのは、IPネットワークの場合、高速道路のトラフィックを保証すればするだけ、ネットワークとしてのリソース(帯域)を一般道から奪っている、という点である。

50 【0004】上記の背景は、もともと「統計多重による帯域の有効活用」を旨とするIPネットワークに、帯域制

御、優先度制御の必要性を生んだ。インターネットにおける差別化思想は、QoS(Quality of Service)、Diffserve、IntServeなどにより、サービスポリシーとして定義、確立されようとしている。差別化思想、サービスポリシーでは、パケット・フロー識別規定(classification)、入力トラフィック特性規定、パケット廃棄優先度規定、帯域保証規定などにより、各々のパケット・フローに対するサービスを管理する。帯域制御方式の代表的な例として、ストリクト・キューイング(Strict Queuing)と、重み付けキューイング(WFQ:Weighted Fair Queueing)がある。ストリクト・キューイングは、優先度が高く設定されたパケットがあれば、必ず優先的に先に送出する単純優先度方式である。図8は単純優先度方式の説明図であり、1<sub>1</sub>~1<sub>4</sub>はクライアント、2はルータ、3はアプリケーションサーバであり、各クライアント1<sub>1</sub>~1<sub>4</sub>がルータ2を介してサーバ3にパケットを送信する場合であり、Pは優先度を表し、値が大きいほど優先度が高くなっている。ルータ2は同時に各クライアント1<sub>1</sub>~1<sub>4</sub>からパケット①~④が各ポートに入力すると優先度が高い順に、図3では③→②→④→①の順にサーバ3に送出する。

【0005】図8は単純な場合であるが、一般的に優先度制御を行うと、単純なFifo(First In First Out)の順序でパケットは送出されなくなる。図9では、①のパケットが⑤のパケットより先に到着しているにも関わらず、⑤が先に送出されている事がわかる。以上のように、ストリクト・キューイングでは、優先度の高いパケットが連続して到着した場合、優先度の低いパケットは、優先度の高いパケットが送出し終わるまで、送出されることはない。このため、優先度の低いパケットはキューイング・メモリにて滞留する可能性が高く、スレッシュホールドを超えて入力があった場合パケット廃棄となる。ストリクト・キューイング方式の特徴としては、非常に厳格な優先度制御が行われ、スケジューリング処理も簡単、単純であるが、各フロー間に自由度のある帯域設定、優先度設定を行うことができない。

【0006】重み付けキューイング方式は、優先度順位をアナログ的なパラメータにて、優先比率(重み付け:weight)を決め、決められた比率に従ってパケットを送出する方式である。図10及び図11は重み付けキューイング方式の説明図であり、サーバ3への一つの出力ポートに対して、クライアント1<sub>1</sub>=12.5%、クライアント1<sub>2</sub>=12.5%、クライアント1<sub>3</sub>=25%、クライアント1<sub>4</sub>=50%、という優先比率が割り当てられている。重み付けキューイング方式では、ストリクト・キューイングのように特定の優先度に厳格に偏ることはなく、この比率に従った帯域割り当て処理が行われる(図11参照)。しかし、重み付けキューイング方式は可変長パケットにおける優先比率制御であり、パケット転送数で重み付け制御すると、最小パケットと最大パケットのサイズの違いが

認識されない。Ethernetフレームを例にとると最小64バイト、最大1522バイトなので、ルータ2において図12に示すようにクライアント1<sub>1</sub>、1<sub>2</sub>からのパケットを交互にサーバ3に送出するものとする、優先比率の誤差が最大24倍近い開きが生じてしまう。この解決方法としては、図13に示すように可変長パケットを固定長サイズに分割して、固定長データの転送数にて制御する方法がある。

【0007】しかし、この方式では、ルータ2の図示しないルーチング部あるいはスイッチ部の前段にsegment処理部(分割処理部)、後段にre-assembly処理部(組立て処理部)が必要となる問題がある。特に、後段での組立て処理部において同時に複数のパケットが組み立て中となる事を想定しなければならず、その各々のパケット毎に論理的、もしくは物理的に個別な組立てバッファを用意し、制御する必要が生じる問題がある。図14はルータのパケット送出制御説明図であり、2aはパケットスイッチ部、2bはスケジューラ、2c<sub>1</sub>~2c<sub>4</sub>は各クライアントからのパケットを組み立てる組立てバッファ、2dは各バッファからパケットを読み出して送出するパケット読み出部である。

【0008】

【発明が解決しようとする課題】(1) 第1の課題  
重み付けキューイング(Weighted Fair Queueing)方式はパケット単位にVirtual Clockという概念を導入する。このVirtual Clockとは、パケットが装置から出力されるまでの時間、すなわち、装置内にパケットが存在する時間である。Virtual Clock情報は、装置内の個々のパケット毎に個別に管理しなければならず、その生成、管理は膨大な処理となる。換言すれば、重み付けキューイング方式におけるスケジューラは、装置内部に存在する全パケットに対して大量かつ複雑な制御情報を生成、管理しなければならず、スケジューリング処理が複雑になり、しかも処理時間が長くなる問題がある。

【0009】(2) 第2の課題

可変長パケットをATMセルのような固定長データに分割し、スケジューラも固定長データ単位にスケジュール管理を行うように作り込むと、スケジューラの制御は比較的簡単にはなる。しかし、そのためには、ルーチング部、スイッチ部の前段に可変長パケットの分割処理が必要となり、かつスイッチ部の後段では組み立て処理が必要となる。そうすると、後段での可変長パケット・組み立てバッファでは、同時に複数のパケットが組み立て中になる事を想定しなければならず、その各々のパケット毎に論理的、もしくは物理的に個別な組み立てバッファを用意する必要があり、装置規模が大きくなる問題がある。

【0010】(3) 第3の課題

物理帯域を100%有効に使用できることが望ましい。このため、出力先に可変長パケットを無駄な空き帯域無し

に、しかも時間的に隙間なくバケット・データを詰めて出力する必要がある。しかし、従来のように複雑なスケジューラ構成では、1回のスケジューリング処理時間が長くなり、出力先へ隙間なくバケット・データを出力することができない。

#### 【0011】(4) 第4の課題

従来のスケジューラは、スケジューリング対象となる待ち行列 (Queue) を1段階で処理しており、優先度設定も全ての行列構成要素 (例えばバケット) を対象として設定されている。この従来方式では、転送許可 (grant) を各行列構成要素間にて巡回させるために Round-Robin 方式を採用している。Round-Robin 方式は、ある待ち行列構成要素が転送許可 (grant) を獲得したら、次のスケジューリング処理において該待ち行列構成要素の優先度を最も低くする方式である。Round-Robin 方式は公正 (fair) なスケジューリングには向いているが、非線形な帯域割り当て制御 (つまり、差別化サービス) にはそのままでは採用できない。例えば、帯域を高めめに設定されている行列構成要素に対しては、転送許可の巡回頻度を上げる必要があるが従来の Round-Robin 方式のスケジューラはかかるスケジューリング制御ができない。

#### 【0012】(5) 第5の課題

可変長バケットを取り扱う物理帯域保証型スケジューラは、大きく2種類に分類される。一つは可変長バケット・データを固定長データに分割して行う方法、もう一つは可変長バケット・データのままだで行う方法である。本発明では、可変長バケット・データのままだで行う方法を採用する。かかる方法において、スケジューラからの転送許可 (grant) は、バケット単位で出力されるのが一般的である。この場合、1回の転送許可において転送されるバケット・データ長は一定ではない。つまり、各行列構成要素において等しく1回の転送許可を巡回したとしても、同じだけの物理帯域を占有したことにはならない。これが数十回、数百回となると、最悪同じ回数だけ転送許可を巡回させているにも関わらず、行列構成要素への割り当て帯域に不公平 (unfair) が生じてしまう。これが、可変長バケットを扱うスケジューラにおいて、帯域保証制御を難しくしている原因である。

#### 【0013】(6) 第6の課題

上記の第5の課題は、1回の転送許可における伝送保証帯域を規定する事により、可変長バケット伝送における不公平感を無くせる。この場合、伝送保証帯域に達する前に送出すべきバケットが存在しなくなった場合の処理を規定する必要がある。かかる処理方式としては、バケットが来るまで待つ方式がある。これは、伝送保証帯域に達するまで、対象行列構成要素に連続して転送許可 (grant) を与え続ける方式であるが、物理帯域の100%利用は不可能であり、無駄な空き帯域が生じる問題がある。

#### 【0014】(7) 第7の課題

前述しているとおり、可変長バケット・スケジューラ

は、公平な帯域割り当て制御を実現するためにバケット転送回数情報を使用することはできない。なぜならば、同じ10個のバケット・データでも一方の行列構成要素からは最小サイズのバケットが多く、もう一方の行列構成要素からは最大サイズのバケットが多い場合、同じ帯域設定としているのに不公平 (unfair) が生じるからである。従って、可変長バケットであっても厳密な帯域制御が行えるようにする必要がある。

【0015】以上より、本発明の目的はストリクトキューイング方式及び重み付けキューイング方式を併用して可変長バケットを分割することなく帯域制御を可能にし、スケジューラ回路規模の縮小化、高速処理化を図ることである。本発明の別の目的は、入力ポートと該ポートより入力するバケットに付加される品質クラスとの組み合わせを待ち行列構成要素とし、バケット単位でなくこの待ち行列構成要素単位で帯域制御、優先度制御することによりスケジューラ回路規模の縮小化、高速処理化を図ることである。本発明の別の目的は、ルーチング部やスイッチ前段のバケット分割処理や後段のバケット組立て処理を不要にすることである。本発明の別の目的は、1回のスケジューリング処理時間が短くなるように、すなわち、スケジューリング処理時間が最小バケットの伝送時間以内で収まるようにすることである。

【0016】本発明の別の目的は、待ち行列構成要素を数段階の絶対優先度グループにグルーピングし、グループ内で Round-Robin 方式により各待ち行列構成要素に順番にバケット送出権を与え、これにより、単純な Round-Robin 方式であっても、優先度の高い特定の行列要素に対して転送許可の巡回頻度を上げられるようにすることである。本発明の別の目的は、優先度グループを単純な絶対優先度方式により選択することにより、シンプルなハードウェアで高速スケジューリング処理を可能にすることである。本発明の別の目的は、各優先度グループに属する行列構成要素を任意に設定でき、また、その数に制限を設けないようにでき、しかも、優先度の高いグループの流量と、優先度の低いグループの流量の比率が偏っていても、帯域制御を可能にすることである。

【0017】本発明の別の目的は、各行列構成要素に対する1回の転送許可による伝送保証帯域を設定して各行列構成要素に割り当てる帯域の制御を行えるようにすることである。本発明の別の目的は、連続して送出するバケットが存在しなくなれば、直ちにその行列構成要素に対する転送許可を解除し、他の行列構成要素に転送許可を渡し、共有している物理帯域において無駄な空き帯域をなくし、帯域を100%利用できるようにすることである。本発明の別の目的は、各入力ポートの品質クラス (行列構成要素) 毎に所定時間当りに送出するバケットのデータ送出流量 (帯域) を設定することにより厳密な帯域制御を実現することである。

【0018】

【課題を解決するための手段】本発明の第1のバケット送出装置は、(1) 優先度がそれぞれ異なる複数のグループの各グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にバケット送出権を与える待ち行列制御部、(2) 少なくとも1つの待ち行列構成要素に応じたバケットが送出待ちになっているグループのうち、優先度が最大のグループをバケット送出グループとして決定するバケット送出グループ決定部、を備え、待ち行列制御部は該バケット送出グループの待ち行列においてバケット送出権のある行列構成要素に応じたバケットを送出する。

【0019】本発明の第2のバケット送出装置は、更に、(3) 入力ポートと該ポートより入力するバケットに付加する各品質クラスの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部を備え、前記待ち行列制御部は、待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にバケット送出権を与える。

【0020】本発明の第3のバケット送出装置は、更に、(4) 待ち行列構成要素毎に送出待ちのバケットを記憶するバッファ部、(5) 送出待ちバケットが記憶されているバッファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちバケットが存在するグループを識別し、これらグループのうち優先度が最大のグループをバケット送出グループとして決定する。

【0021】本発明の第4のバケット送出装置は、更に、(6) 行列構成要素毎に1回のバケット送出権で保証するデータ送出量を設定する保証データ量設定部、(7) バケット送出権を与えられた行列構成要素に応じたバケットの実際の送出量を監視する監視部、(8) 実際のデータ送出量が前記保証データ量と等しくなったとき、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部を備え、前記待ち行列制御部は、該制御信号に基づいてバケット送出権を次の行列構成要素に付与する。

【0022】本発明の第5のバケット送出装置は、更に、(9) 待ち行列構成要素毎に設定時間当りに送出するバケットのデータ送出流量を設定するデータ送出流量設定部、(10) 待ち行列構成要素毎に前記設定時間当りのバケットの実際のデータ送出流量を監視する手段、(11) 待ち行列構成要素毎にバケットの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたバケットの送出を禁止する送出禁止信号を発生するバケット送出禁止手段を備えている。

【0023】上記本発明のバケット送出装置は、どの優

先度グループに転送許可(grant)を出すかという処理と、その優先度グループに属するどの行列構成要素に転送許可(grant)を出力するかという処理との2段構成でバケットを送出するため、バケット送出制御を簡単な構成で行うことができる。又、本発明のバケット送出装置は、可変長バケットデータを固定長に分割することなく、可変長のままでスケジューリング処理を行い、各行列構成要素毎に、すなわち、各入力ポートの品質クラス毎に正確かつ、自由に帯域割り当て制御を実現できる。

【0024】又、本発明のバケット送出装置は、バケット単位でなく、各入力ポートの品質クラス単位でスケジューラ部へ転送要求信号(request)を生成することにより、大幅な回路規模の縮小化、高速化が可能である。又、本発明のバケット送出装置は、固定長に分割されたデータ単位に対して転送許可(grant)するのではなく、あくまで可変長バケット・データ単位に転送許可するため、バケット・データ自体を分割すること無く転送でき、可変長バケットの分割処理(segment処理)、組立て処理(re-assembly処理)を必要としない。又、本発明のバケット送出装置は、スケジューラをハードウェアで構成しているため、高速のスケジューリング判定を可能とし、1回のスケジューラ判定時間を最小バケット長の物理的な伝送時間以内に押えることが可能となる。これにより、スケジューラ処理によるディレイを少なくし、最小バケット・サイズが連続して入力された場合でも、無駄な物理的空き帯域を無くして連続的にバケットの送出が可能である。

【0025】又、本発明のバケット送出装置は、優先度グループを自由に設定できるようにしたから、優先度の高い行列構成要素数と低い優先度の行列構成要素数の比率が偏っていても対応可能である。又、本発明のバケット送出装置は、転送許可された行列構成要素に応じたバケットは伝送保証帯域を超えるまで転送許可されるため、packet-by-packet単位の装置であるにも関わらず、一回の転送許可において物理的な帯域保証(byte/sec、もしくはbit/sec)が可能である。又、実際の帯域が伝送保証帯域となる前にすべてのバケットの送出が完了すれば直ちに転送許可を解除することにより、無駄な物理的空き帯域が発生するのを防止できる。又、本発明のバケット送出装置は、基本的に単純優先度制御のRound-Robinスケジューラであるが、バケット毎の流量情報をフィードバックする構成を採用することにより、ハードウェアによる高速スケジューリング処理が可能となり、しかも、byte/sec、もしくはbit/secという単位で絶対物理帯域の割り当て制御が可能となる。

【0026】

【発明の実施の形態】(A)第1実施例

(a)バケット送出装置の全体の構成

図1は本発明のバケット送出装置の構成図である。このバケット送出装置は、バケットを伝送路に送出する際、

該バケットが到来した入力ポート  $P_0 \sim P_n$  とその品質クラスとに基づいて所定のスケジューリングに従って送出するものである図1において、10はn個の入力ポートと1個の出力ポート有するバケット送出装置、11<sub>0</sub>～11<sub>n</sub>は入力伝送路、12は出力伝送路、 $P_0 \sim P_n$  は入力ポートである。バケット送出装置10は入力ポート  $P_0 \sim P_n$  と該入力ポートから入力するバケットの品質クラス(サービスクラス)との組み合わせに基づいて優先制御及び帯域制御を行いながらバケットを出力伝送路に送出する。

【0027】バケット送出装置10は、SONETの場合であれば各伝送路から入力する伝送路信号を光電変換すると共に、SONET終端処理、バケット分離処理などのインタフェース制御を行う入力側IF部21<sub>0</sub>～21<sub>n</sub>、入力側IF部から入力するバケットに装置内ヘッダの付加などの処理を施して出力するフォワーディング部22<sub>0</sub>～22<sub>n</sub>、フォワーディング部から入力するバケットをQoS毎に振り分けてキューイングするingress側キューイング部23<sub>0</sub>～23<sub>n</sub>、スケジューリング制御に基づいてキューイング部23<sub>0</sub>～23<sub>n</sub>の各キューイングバッファに記憶されているバケットを送出するスケジューリング部24、スケジューリング部から出力するバケットをキューイングして順番に出力するegress側キューイング部25、電光変換、装置内ヘッダの削除、SONETフレーム信号の生成等を行って出力伝送路に送出する出力IF部26を備えている。

【0028】フォワーディング部22<sub>0</sub>～22<sub>n</sub>は図2に示すようにIPバケットPKTに装置内ヘッダFIH(Frame Information Header)を付加して出力するもので、装置内ヘッダFIHは、たとえば、①ルーチングフラグ/マルチキャストフラグ/廃棄フラグ、②受信ポート番号、③出力ポート番号、④QoS(Quority of Service)、⑤バケット長、⑥マルチキャストの出力ポート等を含んでいる。このうちQoSには例えば以下の4つのサービスクラス(品質クラス)、

①iGS(Guranteed Serviceクラス)、

②iCLS(Controlled Load Serviceクラス)、

③iBE(Best Effortクラス)、

④iSpare(予備クラス)、がある。iGSは要求された帯域を保証するサービスクラス、iCLSは所定の最低帯域を保証するサービスクラス、iBEは品質規定がないサービスクラス、iSpareは予備クラスである。各フォワーディング部22<sub>0</sub>～22<sub>n</sub>は、IPアドレス(発信元アドレス/宛先アドレス)、受信ポート番号、レイヤ4プロトコル(TCP/UDP)、サービス種別(WEB/FTP/TELNET)のいずれか、あるいはこれらの組み合わせに基づいてサービスクラスを決定してバケットに付加する。

【0029】ingress側キューイング部23<sub>0</sub>～23<sub>n</sub>はそれぞれサービスクラス毎に総計4つのキューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>( $i=0 \sim n$ )を備え、対応するフォワー

ディング部22<sub>0</sub>～22<sub>n</sub>から入力するバケットのQoSを参照し、サービスクラス毎にバケットをキューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>に記憶する。又、ingress側キューイング部は、スケジューリング部24からの転送許可により順番にバケットを送出する。

【0030】(b)スケジューリング部

スケジューリング部24はキューイング部23<sub>0</sub>～23<sub>n</sub>の各キューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>...QBF<sub>0,1</sub>～QBF<sub>3,4</sub>にキューイングされているバケットをスケジューリング制御により順番に送出するもので、図3、図4に示す構成を備えている。尚、図3にはingress側キューイング部のキューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>...QBF<sub>0,1</sub>～QBF<sub>3,4</sub>を含めて示している。図3、図4は図1におけるスケジューリング部24の構成図であり、二分した一方を図3に、他方を図4に示している。スケジューリング部24は大別すると、ポート対応部31<sub>0</sub>～31<sub>n</sub>( $n=7$ )、スケジューラ32、バケット合成部33を有し、各ポート対応部31<sub>0</sub>～31<sub>n</sub>は同一の構成を備えている。

【0031】(b-1)ポート対応部

ポート対応部31<sub>0</sub>は、ポート  $P_0$  から入力するバケットをサービスクラス(品質クラス)毎に振り分けてキューイングし、スケジューラ32からの転送許可に基づいてバケットを送出する部分で、GS用バッファQBF<sub>0,1</sub>、CLS用バッファQBF<sub>0,2</sub>、BE用バッファQBF<sub>0,3</sub>、Spare用バッファQBF<sub>0,4</sub>を備えている。同様に、ポート対応部31<sub>i</sub>( $i=1 \sim 7$ )はポート  $P_i$ ( $i=1 \sim 7$ )から入力するバケットをサービスクラス毎に振り分けてキューイングし、スケジューラ32からの転送許可に基づいてバケットを送出する部分であり、GS用バッファQBF<sub>0,1</sub>、CLS用バッファQBF<sub>0,2</sub>、BE用バッファQBF<sub>0,3</sub>、Spare用バッファQBF<sub>0,4</sub>を備えている。各ポート対応部31<sub>i</sub>( $i=0 \sim 7$ )は上記構成に加えて、リクエスト信号発生部31a、リクエスト禁止部31b、転送許可部(Grant部)31c、連続転送許可部31d、伝送保証帯域記憶部31e<sub>0</sub>～31e<sub>7</sub>、レジスタ31f、設定値ロード部31g、オクテットカウンタ31h、比較部31i、合成部31jを有している。

【0032】リクエスト信号発生部31aは、各キューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>にそれぞれ転送待ちバケットが存在するとき、転送許可を要求するリクエスト信号(Request)R<sub>0,1</sub>～R<sub>3,4</sub>を発生する。リクエスト禁止部31bは、所定サービスクラスのバケットを設定時間内に設定量以上送出したとき、該サービスクラスのバケットの送出を禁止するためにリクエスト信号をマスクし、リクエスト信号がスケジューラ32に入力するのを禁止する。転送許可部(Grant部)31cはスケジューラ32から入力する転送許可信号(Grant)に基づいてキューイングバッファQBF<sub>0,1</sub>～QBF<sub>3,4</sub>にバケットの転送を許可する。連続転送許可部31dは連続してバケットの転送を許可信号を送出すると共に、所定の条件が成立したとき次の行列構成要素にバケット送出権を与えるようスケジュー

10

20

30

40

50



ラ32に巡回信号(Rotate)を入力する。所定の条件とは、①1回の転送許可により送出可能なデータ量(伝送保証データ量)を送出したとき、②実際の送出データ量が1回の転送許可により送出可能な伝送保証データ量に等しくなる前にバケットがバッファからすべて送出されたときであり、いずれかの条件が成立すると、転送許可信号(Grant)の出力を停止すると共に、巡回信号(Rotate)を出力する。

【0033】伝送保証データ量記憶部31e<sub>1</sub>~31e<sub>n</sub>は1回の転送許可により送出可能な伝送保証データ量(byte/sec)TH<sub>1</sub>~TH<sub>n</sub>をサービスクラス毎に記憶し、設定値ロード部31gは転送許可された入力ポートのサービスクラスに応じた設定データ量THiをレジスタ31fにロードする。オクテットカウンタ31hは実際に送出されるバケットデータのオクテット数をカウントし、比較部31iは実際に送出されるバケットのオクテット数とレジスタ31fに設定された伝送保証データ量THiを比較し、一致したときStop信号を出力すると共に、オクテットカウンタ31hをリセットする。合成部31jは各キューイングバッファQBf<sub>0</sub>~QBf<sub>n</sub>から出力するバケットを合成してバケット合成部33を介してegress側キューイング部25に送出する。

#### 【0034】(b-2) スケジューラ

スケジューラ32において、レジスタ32a<sub>01</sub>~32a<sub>n</sub>には入力ポートP<sub>0</sub>~P<sub>n</sub>のサービスクラス(iGS, iCLS, iBE, iSpare)毎に一定時間内に送出可能な最大データ量TH<sub>1</sub>~TH<sub>n</sub>(byte/sec)(i=0~n)が設定される。カウンタ32b<sub>01</sub>~32b<sub>n</sub>は各入力ポートP<sub>0</sub>~P<sub>n</sub>のサービスクラス(iGS, iCLS, iBE, iSpare)毎に設定時間内に実際に送出したバケットデータ量をカウントするものである。フロー識別部32cはバケットに付加された装置内ヘッダFIDを参照して送出バケットの入力ポート番号及びサービスクラスを識別し、識別した入力ポート番号/サービスクラスに応じたカウンタ32b<sub>01</sub>~32b<sub>n</sub>により送出データ量をカウントさせる。比較部32e<sub>01</sub>~32e<sub>n</sub>は、各入力ポートの品質クラス毎のバケットデータ送出量が設定最大データ量TH<sub>1</sub>~TH<sub>n</sub>と等しくなった時、すなわち、カウンタ32b<sub>01</sub>~32b<sub>n</sub>の計数値が設定最大データ量TH<sub>1</sub>~TH<sub>n</sub>に等しくなったとき、設定時間が経過するまで該バケットの送出を禁止する送出禁止信号INHを発生する。リセット部32fは設定時間毎に各カウンタの計数値をリセットする。

【0035】リクエスト集積部32gはポート対応部31<sub>0</sub>~31<sub>n</sub>から出力する32個のリクエスト信号を集積して、順番に並べて出力し、グループ設定部32hは、第1~第4優先グループを設定すると共に、入力ポートと該ポートより入力するバケットのサービスクラスとの組み合わせを待ち行列構成要素とし、これら各行列構成要素が属するグループを設定する。図5(A)の例では、(1) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラス

CSを第1優先度グループの行列構成要素P<sub>0</sub>-CS~P<sub>n</sub>-CSとし、(2) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスCLSを第2優先度グループの行列構成要素P<sub>0</sub>-CLS~P<sub>n</sub>-CLSとし、(3) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスBEを第3優先度グループの行列構成要素P<sub>0</sub>-BE~P<sub>n</sub>-BEとし、(4) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスSpareを第4優先度グループの行列構成要素P<sub>0</sub>-Spare~P<sub>n</sub>-Spareとする。

【0036】図5(B)は、(1) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスCSを第1優先度グループの行列構成要素P<sub>0</sub>-CS~P<sub>n</sub>-CSとし、(2) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスCLS、入力ポートP<sub>0</sub>のサービスクラスCS、入力ポートP<sub>0</sub>のサービスクラスBEをそれぞれ第2優先度グループの行列構成要素P<sub>0</sub>-CS、P<sub>0</sub>-CLS~P<sub>0</sub>-BEとし、(3) 各入力ポートP<sub>1</sub>~P<sub>n</sub>におけるサービスクラスBEを第3優先度グループの行列構成要素P<sub>1</sub>-BE~P<sub>n</sub>-BEとし、(4) 各入力ポートP<sub>0</sub>~P<sub>n</sub>におけるサービスクラスSpareを第4優先度グループの行列構成要素P<sub>0</sub>-Spare~P<sub>n</sub>-Spareとする。以上のように、各優先度グループには適宜必要な行列構成要素を含ませることができる。

【0037】優先度マッピング部32iは32個のリクエスト信号を優先度グループ毎に仕分けして出力する。すなわち、優先度マッピング部32iは32個のリクエスト信号を、(1) 第1優先度グループに属する入力ポート/QoS(行列構成要素)に応じたリクエスト信号、(2) 第2優先度グループに属する行列構成要素に応じたリクエスト信号、(3) 第3優先度グループに属する行列構成要素に応じたリクエスト信号、(4) 第4優先度グループに属する行列構成要素に応じたリクエスト信号のそれぞれに分類して出力する。絶対優先度判定部32jは絶対優先度判定処理を行い、所属する少なくとも1つの行列構成要素から転送要求のリクエスト信号(Request)発生しているグループのうち優先度が最大のグループをバケット送出グループとして決定し、該グループにenable信号(有効信号を)を出力する。すなわち、第1優先グループに所属する行列構成要素からリクエスト信号が発生していれば、該第1優先グループにイネーブル信号を入力する。しかし、第1優先グループに所属する行列構成要素からリクエスト信号が発生せず、第2優先グループに所属する行列構成要素からリクエスト信号が発生していれば、該第2優先グループにイネーブル信号を入力する。

【0038】各優先グループ毎に設けられたラウンドロビン回路33k<sub>i</sub>(i=1~4)は第i優先度グループの各行列構成要素で待ち行列を生成し、各待ち行列構成要素にラウンドロビン方式で順番に転送許可(バケット送出権)を与える。Round-Robin方式は、基本的な巡回方向に順番に、かつ、均等に転送許可(バケット送出権)を行列構成要素に与え、かつ、前回転送許可した行列構成要素を今回では最低優先度となるようにするルールで転送許

可を与える方式である。基本的な巡回方向は、優先度グループ設定時に固定的に設定される。ここで注意する事は、転送許可は、あくまで送出待ちバケットが存在する待ち行列構成要素(リクエスト信号を出力している待ち行列構成要素)に対して出力される事である。

【0039】図6はRound-Robin方式の説明図であり、図5(A)の第1優先グループの待ち行列を示している。斜線で示す位置が行列先頭で、優先度が最も高く、以下順に時計方向に優先度が低くなっており、図6(A)では、行列構成要素 $P_0$ -GSの優先度が最高で、行列構成要素 $P_7$ -GSの優先度が最低になっている。Round-Robin方式では、 $P_0$ -GS→ $P_1$ -GS→ $P_2$ -GS→ $P_3$ -GS→ $P_4$ -GS→ $P_5$ -GS→ $P_6$ -GS→ $P_7$ -GS→ $P_0$ -GS→…の順番に、かつ、均等に各行列構成要素に転送許可(バケット送出権)を与える。すなわち、第1優先グループが図6

(A)に示す状態にあるとき、この第1優先グループに転送許可が出されると(enable="1")、先頭の行列構成要素 $P_0$ -GSにバケット送出権を与え(Grant)、該行列構成要素 $P_0$ -GSに応じた送出待ちバケットを一定量送出させる。又、送出完了により巡回信号(Rotate)が発生すると、図6(B)に示すように、各行列構成要素の位置を反時計方向に1つづシフトし、次の $P_1$ -GSを先頭の行列構成要素とし、 $P_0$ -GSを最低優先度の行列構成要素とする。そして、再び第1優先グループに転送許可が出されると、先頭の行列構成要素 $P_1$ -GSにバケット送出権を与え、該行列構成要素 $P_1$ -GSに応じたバケットを一定量送出させる。以後、同様に順番に、かつ、均等に各行列構成要素に転送許可(送出権)を与える。ただし、先頭の第1優先度の行列構成要素に応じた送出待ちバケットが存在しなければ、すなわち、該行列構成要素から転送要求のリクエスト信号(Request)が発生していなければ、第2優先度の行列構成要素(図6(A)では $P_1$ -GS)にバケット送出権を与える。又、第2優先度の行列構成要素 $P_1$ -GSに応じた送出待ちバケットが存在しなければ、第3優先度の行列構成要素 $P_2$ -GSにバケット送出権を与え、以下同様に送出待ちバケットが存在する行列構成要素にバケット送出権を与える。

【0040】以上より、絶対優先度判定部32jからイネーブル信号(有効信号)が発生すると、該イネーブル信号が入力したランドロビン回路33kiは、自身の待ち行列の先頭行列構成要素からリクエスト信号(Request)が入力されているか調べ、入力されていれば対応するキューイングバッファに対してバケットの転送許可(Grant)を出力する。又、設定されている伝送保証データ量のバケットデータの転送が完了し、あるいは、全待ちバケットの送出が完了し、連続転送許可部31dから巡回信号(Rotate)が入力すれば、先頭の行列要素を末尾に移動させ、残りの行列構成要素を1つづ前進させ、これにより、次の行列構成要素を先頭に位置させる。

【0041】(c)全体の動作

以上より、図3、図4には8ポート( $P_0 \sim P_7$ )のスケジューリング部が示され、サービスクラス単位に待ち行列要素Queue-iGS, Queue-iCLS, Queue-iBE, Queue-iSpareが構成され、各待ち行列構成要素に応じたキューイングバッファQBF<sub>i1</sub>~QBF<sub>i8</sub>が設けられている。スケジューリング部24はこの待ち行列要素に対して帯域制御を行う。各待ち行列要素に応じた送出待ちバケットがキューイングバッファに存在するとき、リクエスト信号発生部31aは転送要求のリクエスト信号(Request)を発生してスケジューラ32に inputs する。絶対優先度判定部32jは所属する行列構成要素のリクエスト信号(Request)が少なくとも1つ入力されているグループのうち優先度が最大のグループをバケット送出グループとして決定し、該グループにenable信号(有効信号)を出力する。

【0042】各優先度グループには、個別のRound-Robin回路33k<sub>1</sub>~33k<sub>8</sub>が用意されておりenable信号が与えられたグループのRound-Robin回路33k<sub>1</sub>が有効となる。このRound-Robin回路33k<sub>1</sub>は、リクエスト信号(Request)を出力している行列構成要素の中から最も優先度の高い行列構成要素( $P_0$ -GSとする)を求め、該行列構成要素 $P_0$ -GSに対して転送許可信号(Grant)を出力する。転送許可信号(Grant)は、行列構成要素 $P_0$ -GSに応じた入力ポート $P_0$ の該当するキューイングバッファQBF<sub>01</sub>にGrant回路31cを介して伝えられる。これにより、キューイングバッファQBF<sub>01</sub>は記憶するバケットの転送を開始する。この時、行列構成要素 $P_0$ -GSに対して予め設定されている伝送保証帯域値(伝送保証データ量byte/sec)TH<sub>1</sub>をレジスタ31fにロードする。これは、行列構成要素 $P_0$ -GSに対して一回の転送許可において許されている物理帯域を守るためである。基本的に行列構成要素 $P_0$ -GSに応じたキューイングバッファQBF<sub>01</sub>からのバケット読み出しは、実際の送出データ量が伝送保証データ量TH<sub>1</sub>と等しくなるまで連続して行われる。すなわち、比較部31iにより実際の送出データ量が保証データ量に達したことが検出されると、連続Grant回路31dは巡回信号(Rotate)をスケジューラ部32へ出力する。又、連続Grant回路31dは、実際の送出データ量が伝送保証データ量に達しなくても、待ちバケットをすべて送出してバッファが空になったことを検出すると転送許可状態(Grant状態)を放棄し、巡回信号(Rotate)をスケジューラ部32へ出力する。

【0043】キューイングバッファQBF<sub>01</sub>から読み出されたバケットはフロー識別部32cにより入力ポート番号/サービスクラスが識別され、行列構成要素は $P_0$ -GSのバケットフローであると識別される。これにより、スケジューラ部32における行列構成要素 $P_0$ -GSに応じたカウンタ32b<sub>01</sub>はバケットの流量をカウントする。各カウンタ32b<sub>01</sub>~32b<sub>08</sub>は同様にバケットデータの流量をカウントし、設定時間内における対応する行列構成要素に応じたバケットの流量を計測し、該設定時間が

くれば自動的にクリアされる。各行列構成要素には予め設定時間における最大帯域（最大送出データ量byte/sec）が設定されており、この時間内にカウンタ32b<sub>0,1</sub>～32b<sub>7,4</sub>の計数値が設定値と等しくなると、比較部32e<sub>0,1</sub>～32e<sub>7,4</sub>は送出禁止信号INHを出力し、リクエスト禁止部31bはこの送出禁止信号により対応する行列構成要素に応じたリクエスト信号（Request）をマスクする。この結果、いくら行列構成要素に応じた送出待ちパケットが存在していても、スケジューラ部32からの転送許可は出力されなくなる。

【0044】本発明では優先度グループの選択論理を単純な絶対優先度方式としている。単純な絶対優先度方式で良い理由を説明する。ここでいう単純な絶対優先度方式とは、「あるスケジューリング処理において、その瞬間の最も優先度の高いグループが必ず選択される」というものである。つまり、高優先度グループにある程度パケットが存在している場合、転送許可（Grant）はその高優先度グループに属する行列構成要素間にて巡回し、それより低優先度グループへ転送許可が渡される事はない。従って、この絶対優先度方式だけでは、最悪のケースの場合、低優先度グループの行列構成要素にパケット送出権が与えられない可能性が出てくる。しかし、本発明では、各行列構成要素単位に流量カウンタ32b<sub>0,1</sub>～32b<sub>7,4</sub>が用意されており、常に設定流量と実際の流量との比較処理を行っており、実際の流量が設定流量を超えた事が検出されると、その行列構成要素からの転送要求のリクエスト信号（Request）をマスクする。このため、設定流量に達した行列構成要素に対して以後設定時間内にスケジューラから転送許可（Grant）が出力されることはない。これにより、本発明によれば、各行列構成要素毎に実際の流量を監視して不当な帯域割り当てを防止できる。

【0045】本発明では、1回の転送許可による伝送保証データ量に実際の送出データ量が達しなくても、待ちパケットをすべて送出してバッファが空になったことを検出すると転送許可状態を解除し、他の行列構成要素に対して転送許可（パケット送出権）を渡すようにしている。このようにすれば、共有している物理帯域において無駄な空き帯域なくなり、帯域を100%利用する事が可能となる。しかし、この方式は、伝送保証帯域に達しないにも関わらず、転送許可状態を解除することを許した事になり、各行列構成要素に割り当てた帯域保証が守られるのかという疑問が生まれる。しかし、本発明では設定時間毎に前述のマスク制御により帯域を制御するため何らの問題は生じない。

#### 【0046】（B）第2実施例

図7は本発明の第2実施例であるn組の入出力ポートを有するパケットスイッチ装置の全体構成図である。このパケットスイッチ装置は、パケットを所定の伝送路に送出する際、ポート部毎に該パケットが到来した入力ポ

ートとパケットのサービスクラスとに基づいて所定のスケジューリングを行って伝送路に送出する。図7において、51<sub>0</sub>～51<sub>n</sub>は1組の入出力ポートに対応して設けられたポート部、52は各入力ポートから到来するパケットを所定の出力ポートにスイッチングするポート接続部（スイッチ部）であり、各ポート部51<sub>0</sub>～51<sub>n</sub>は同一の構成を備えている。各ポート部51<sub>0</sub>～51<sub>n</sub>は、パケット入力方向に関して、入力側IF部（図示せず）から入力するパケットに所定の処理を施して出力するフォワーディング部61、フォワーディング部から入力するパケットを待ち行列（キュー）に接続して順番に出力するingress側キューイング部62を備え、出力方向に関してegress側キューイング部63、所定のスケジューリングに従ってパケットを出力側IF部から伝送路に送出するスケジューリング部64を備えている。

【0047】フォワーディング部61は第1実施例のフォワーディング部22<sub>0</sub>～22<sub>n</sub>と同様にIPパケットに装置内ヘッダFIH（Frame Information Header）を付加して出力する。ingress側キューイング部62はフォワーディング部から入力するパケットを待ち行列（キュー）に接続して順番に出力するもので、例えば1個のキューイングバッファを備えFIFO形式でパケットを蓄積/出力するもの、あるいは、QoSのサービスクラス毎にキューイングバッファを備え、各バッファにサービスクラス対応のパケットを蓄積し、優先度の高いサービスクラスのパケットから出力するもの等がある。スイッチ部52はn入力、n出力のスイッチであり、各ポート部51<sub>0</sub>～51<sub>1</sub>、（n=7）から入力するパケットをスイッチングして所定のポート部に出力する。egress側キューイング部63は多数のキューイングバッファを備え、スイッチ部52から出力するパケットを該パケットの入力ポート毎に仕分けると共にQoSのサービスクラス毎に仕分けしてそれぞれを各キューイングバッファに記憶する。入力ポート数を8、QoSのサービスクラス数を4（iGS、iCLS、iBE、iSpare）とすれば、egress側キューイング部63は各入力ポート毎に4個づつ総計32個のキューイングバッファQB F<sub>0,1</sub>、QB F<sub>0,2</sub>、...QB F<sub>7,4</sub>を備え、それぞれにパケットを記憶する。

【0048】スケジューリング部64はegress側キューイング部63の各バッファQB F<sub>0,1</sub>～QB F<sub>7,4</sub>にキューイングされているパケットを所定のスケジューリングに従って送出する。すなわち、スケジューリング部64は、各バッファに蓄積されたパケットに対して、帯域を管理しながらパケット単位に転送許可（grant）を出力し、パケットを出力側IF部を介して伝送路に送出する。スケジューリング部64は第1実施例と同様に図3、図4に示す構成を備えている。ただし、キューイングバッファQB F<sub>0,1</sub>～QB F<sub>7,4</sub>を含めて示している。ここで注意すべき点は、入力ポートP<sub>0</sub>から入力したパケットはサービスクラス（iGS、iCLS、iBE、iSpare）毎に仕分けされてキューイングバ

17

ファQBF<sub>0,1</sub>～QBF<sub>0,4</sub>にキューイングされ、入力ポートP<sub>1</sub>から入力したパケットはサービスクラス(iGS, iCLS, iBE, iSpare)毎に仕分けされてキューイングバッファQBF<sub>1,1</sub>～QBF<sub>1,4</sub>にキューイングされ、以下同様に入力ポートP<sub>7</sub>から入力したパケットはサービスクラス(iGS, iCLS, iBE, iSpare)毎に仕分けされてキューイングバッファQBF<sub>7,1</sub>～QBF<sub>7,4</sub>にキューイングされる点である。スケジューリング部64の動作は第1実施例と同じである。以上では、本発明をパケットスイッチング装置に適用した場合について説明したが、パケットルーチング装置にも本発明を適用することができる。

#### 【0049】・付記

(付記1) 優先度がそれぞれ異なる複数のグループに属するパケットを順に送出するパケット送出装置において、グループ毎に待ち行列を生成し、各待ち行列を構成する要素に順番にパケット送出権を与える待ち行列制御部、少なくとも1つの待ち行列構成要素に応じたパケットが送出待ちになっているグループのうち、優先度が最大のグループをパケット送出グループとして決定するパケット送出グループ決定部、を備え、待ち行列制御部は前記決定されたパケット送出グループの待ち行列においてパケット送出権のある行列構成要素に応じたパケットを送出することを特徴とするパケット送出装置。

【0050】(付記2) 前記待ち行列制御部は、パケット送出権のある行列構成要素に応じたパケットが送出待ちになっていなければ次の順位の行列構成要素に応じた送出待ちパケットを送出する、ことを特徴とする付記1記載のパケット送出装置。

(付記3) 入力ポートと該ポートより入力するパケットに付加する各品質クラスの組み合わせを待ち行列構成要素とし、それぞれの待ち行列構成要素が属するグループを設定するグループ設定部、を備え、前記待ち行列制御部は、待ち行列毎に各待ち行列構成要素にラウンドロビン形式で順番に、かつ、均等にパケット送出権を与える、ことを特徴とする付記1記載のパケット送出装置。

【0051】(付記4) 待ち行列構成要素毎に送出待ちのパケットを記憶するバッファ部、送出待ちパケットが記憶されているバッファ部に応じた待ち行列構成要素毎に送出要求のリクエスト信号を発生するリクエスト発生部、を備え、前記送出グループ決定部は、各グループに属する待ち行列構成要素からのリクエスト信号の有無に基づいて送出待ちパケットが存在するグループを識別し、これらグループのうち優先度が最大のグループをパケット送出グループとして決定する、ことを特徴とする付記1記載のパケット送出装置。

【0052】(付記5) 行列構成要素毎に1回のパケット送出権で保証するデータ送出量を設定する保証データ量設定部、パケット送出権を与えられた行列構成要素に応じたパケットの実際の送出量を監視する監視部、実際のデータ送出量が前記保証データ量と等しくなったと

18

き、送出権を次の行列構成要素に渡すための制御信号を出力する制御信号発生部、を備え、前記待ち行列制御部は、該制御信号に基づいてパケット送出権を次の行列構成要素に付与する、ことを特徴とする付記1または3記載のパケット送出装置。

(付記6) 前記制御信号発生部は、実際の帯域が保証帯域を越える前にパケットがバッファからすべて送出されたことを検出したとき、前記制御信号を出力する、ことを特徴とする付記5記載のパケット送出装置。

【0053】(付記7) 待ち行列構成要素毎に設定時間当りに送出するパケットのデータ送出流量を設定するデータ送出流量設定部、待ち行列構成要素毎に前記設定時間当りのパケットの実際のデータ送出流量を監視する手段、待ち行列構成要素毎にパケットの実際のデータ送出流量を監視し、該データ送出流量が前記設定データ送出流量に等しくなった時、前記設定時間が経過するまで該待ち行列構成要素に応じたパケットの送出を禁止する送出禁止信号を発生するパケット送出禁止手段、を備えたことを特徴とする付記1又は3記載のパケット送出装置。

【0054】(付記8) パケット送出禁止手段は、前記設定時間毎に実際のデータ送出流量を零にクリアして送出禁止を解除すること、を特徴とする付記7記載のパケット送出装置。

(付記9) 前記リクエスト発生部は、所定の待ち行列構成要素に対して送出禁止信号が発生したとき、該待ち行列構成要素に応じた送出待ちパケットが存在してもリクエスト信号を発生しない、ことを特徴とする付記8記載のパケット送出装置。

#### 【0055】

【発明の効果】以上本発明によれば、ストリクトキューイング方式及び重み付けキューイング方式を併用して可変長パケットを分割することなく帯域制御を可能にし、スケジューラの回路規模の縮小化、高速処理化を図ることができる。又、本発明によれば、入力ポートと該入力ポートより入力するパケットの品質クラスの組み合わせを待ち行列構成要素とし、パケット単位でなくこの待ち行列構成要素単位で帯域制御、優先度制御するようにしたから、スケジューラの回路規模の縮小化、高速処理化を図ることができる。又、本発明によれば、パケット分割処理やパケット組立て処理を不要にすることができる。

【0056】又、本発明によれば、1回のスケジューリング処理時間が短くなるように、すなわち、スケジューリング処理時間が最小パケットの伝送時間以内で収まるように構成できる。このため、たとえ最小パケットを連続して出力する状況においても、空き帯域なくパケット・データを連続して出力することができる。又、本発明によれば、待ち行列構成要素を数段階の絶対優先度グループにグルーピングし、グループ内でRound-Robin方式により各待ち行列構成要素に順番にパケット送出権を与

えるようにしたから、単純なRound-Robin方式であっても、優先度の高い特定の行列要素に対して、転送許可の巡回頻度を上げることが可能となる。又、本発明によれば、優先度グループを単純な絶対優先度方式により選択することにより、簡単なハードウェアによる高速スケジューリング処理ができる。

【0057】又、本発明によれば、各優先度グループに属する行列構成要素を任意に設定でき、また、その数に制限を設けないようにできる。しかも、優先度の高いグループの流量と、優先度の低いグループの流量の比率が偏っていても、帯域制御が可能である。又、本発明によれば、行列構成要素毎に1回の転送許可による伝送保証帯域を設定するようにしたから、各行列構成要素に割り当てる帯域の制御及び帯域の管理が可能となる。又、本発明によれば、送出する待ちバケットが存在しなくなれば、直ちにその行列構成要素に対する転送許可を解除し、他の行列構成要素に転送許可を渡すようにしたから、共有している物理帯域において無駄な空き帯域をなくすることができる。

【0058】又、本発明によれば、各行列構成要素毎に所定時間当りに送出するバケットのデータ送出流量（帯域）を設定し、各行列構成要素毎にバケットのデータ送出流量を監視し、該データ送出流量が設定データ送出流量を越えた時、所定時間が経過するまで行列構成要素に応じたバケットの送出を禁止するようにしたから、厳密な帯域制御を実現できる。

【図面の簡単な説明】

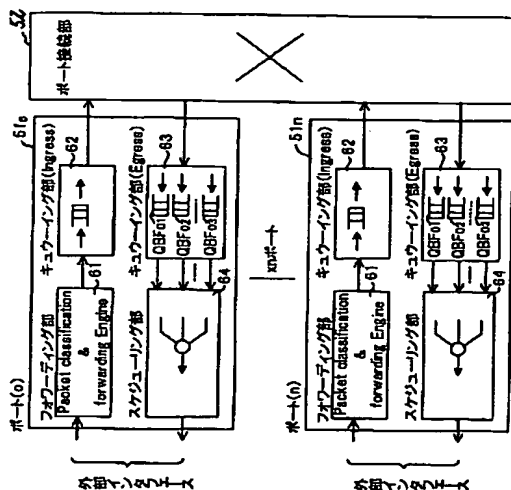
【図1】本発明のバケット送出装置の構成図である。

【図2】装置内ヘッダの構成図である。

【図3】スケジューリング部の詳細構成図（その1）で\*

【図7】

本発明のバケットスイッチ装置の全体の構成



\*ある。

【図4】スケジューリング部の詳細構成図（その2）である。

【図5】優先度グループの例である。

【図6】ラウンドロビン方式の説明図である。

【図7】本発明のバケットスイッチ装置の全体構成図である。

【図8】単純優先度方式の説明図である。

【図9】単純優先度方式の別の説明図である。

10 【図10】重み付けキューイング方式の説明図である。

【図11】重み付けキューイング方式の別の説明図である。

【図12】重み付けキューイング方式の問題点説明図である。

【図13】重み付けキューイング方式の問題点解決手法の説明図である。

【図14】重み付けキューイング方式の解決手法の問題点説明図である。

【符号の説明】

10・・・バケット送出装置

11<sub>0</sub>～11<sub>n</sub>・・・入力伝送路

12・・・出力伝送路

21<sub>0</sub>～21<sub>n</sub>・・・入力側IF部

22<sub>0</sub>～22<sub>n</sub>・・・フォワーディング部

23<sub>0</sub>～23<sub>n</sub>・・・ingress側キューイング部

24・・・スケジューリング部

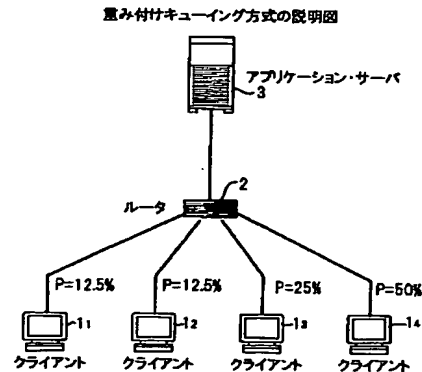
25・・・egress側キューイング部

26・・・出力IF部

P<sub>0</sub>～P<sub>n</sub>・・・入力ポート

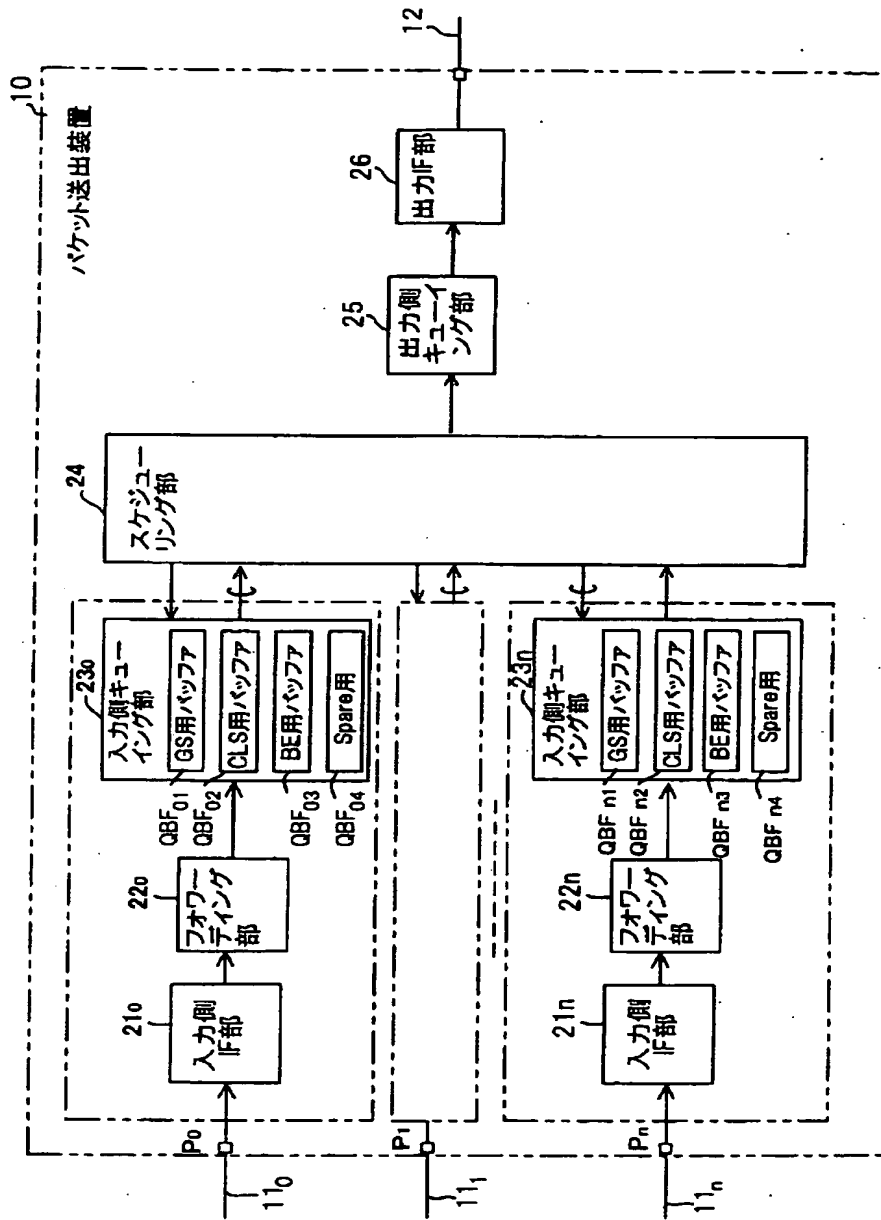
QB<sub>F1</sub>～QB<sub>F<sub>i</sub></sub> (i=0～n)・・・キューイングバッファ

【図10】



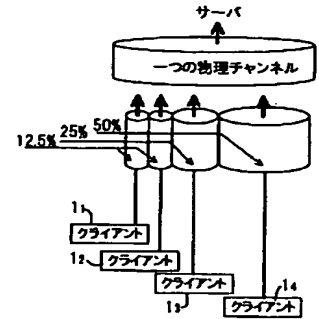
【図1】

## 本発明の packets 送出装置の構成

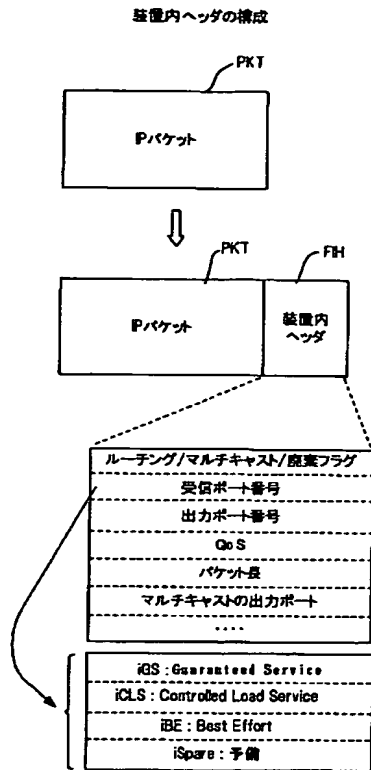


【図11】

## 重み付けキューイング方式の別の説明図



【図2】



【図5】

優先度グループの例

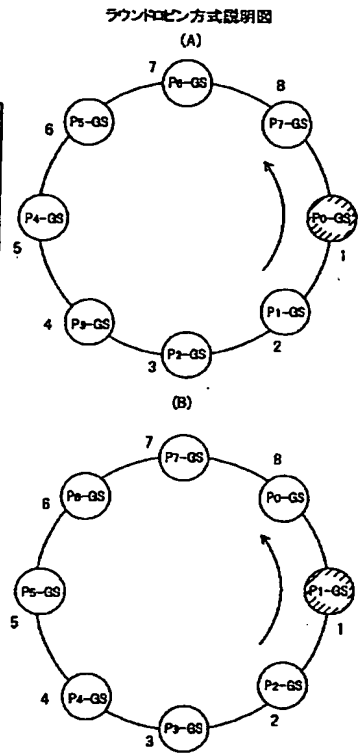
(A)

優先順位グループ	ポートNo. & QoS	個数
第1優先グループ	Po-GS ~ P7-GS	8
第2優先グループ	Po-CLS ~ P7-CLS	8
第3優先グループ	Po-BE ~ P7-BE	8
第4優先グループ	Po-Spare ~ P7-Spare	8

(B)

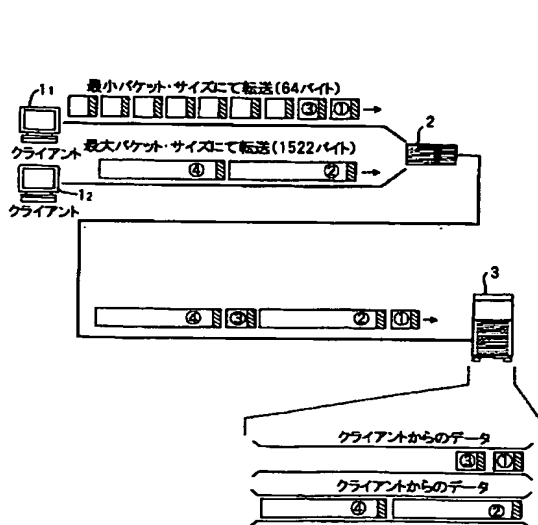
優先順位グループ	ポートNo. & QoS	個数
第1優先グループ	Po-GS ~ P6-GS	7
第2優先グループ	P7-GS Po-CLS ~ P7-CLS Po-BE	10
第3優先グループ	P1-BE ~ P7-BE	7
第4優先グループ	Po-Spare ~ P7-Spare	8

【図6】



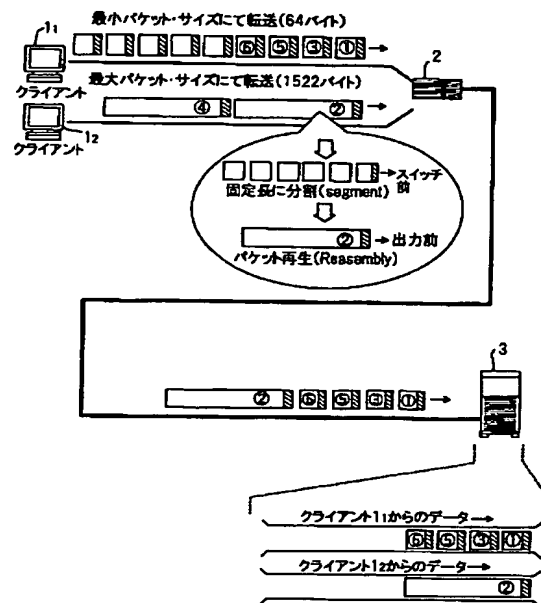
【図12】

重み付けキューイング方式の問題点説明図



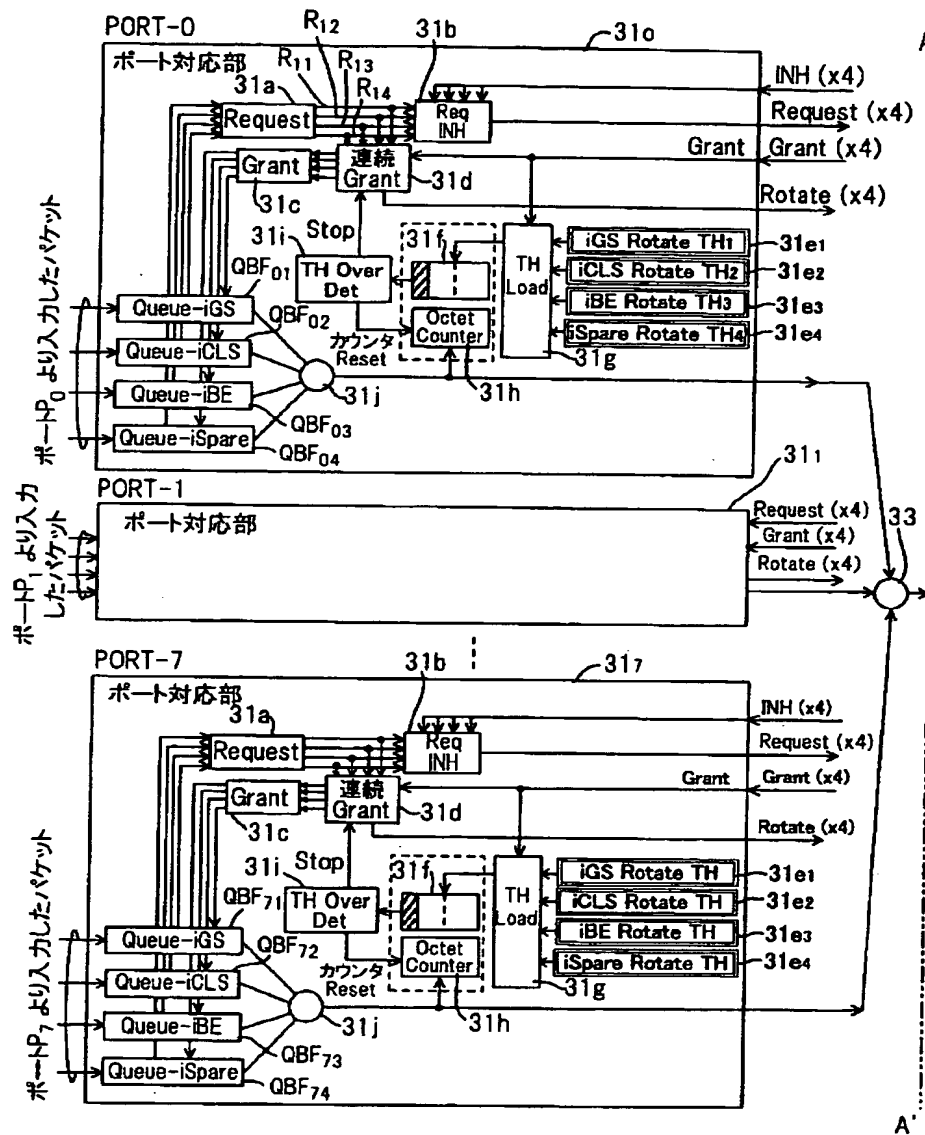
【図13】

重み付けキューイング方式の問題点解決手法の説明図



【図3】

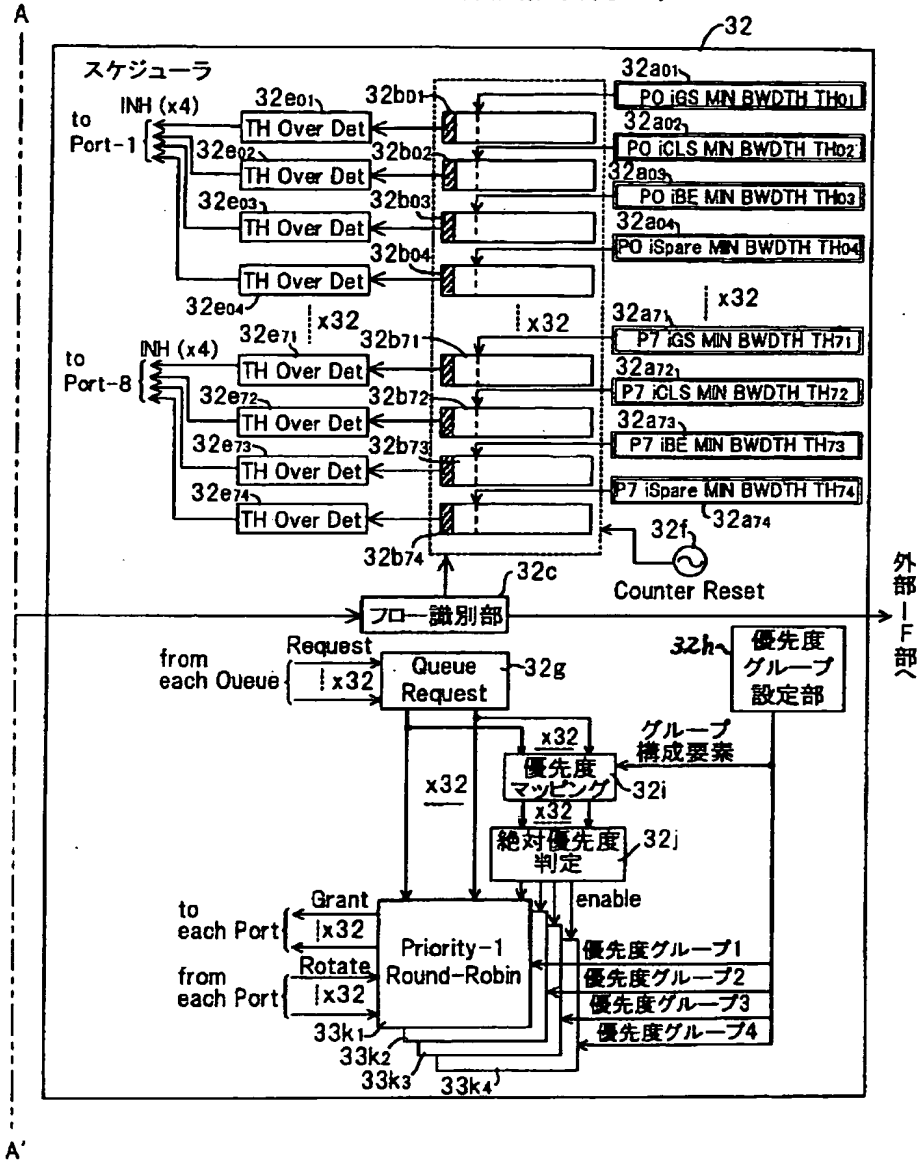
スケジューリング部の詳細構成図(その1)





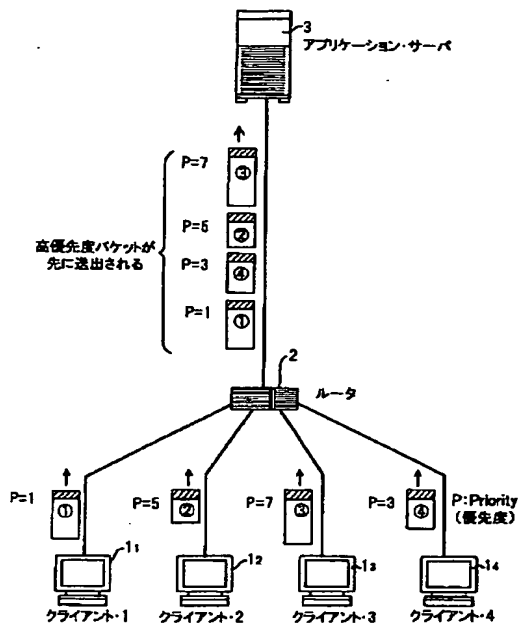
【図4】

スケジューリング部の詳細構成図(その2)



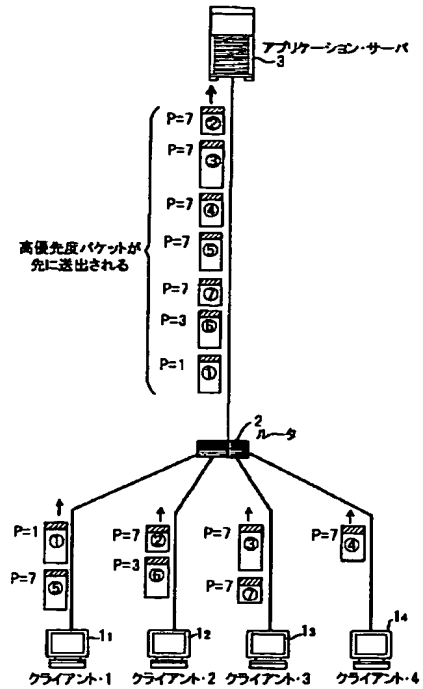
【図8】

単純優先度方式説明図



【図9】

単純優先度方式の別の説明図





**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**